Universität Hamburg
Fachbereich Informatik
Vogt-Kölln-Straße 30
22527 Hamburg

Seminararbeit

Epistemische Logiken

Seminararbeit im Rahmen der Veranstaltung „Logiken in Multiagentensystemen“ (LV-Nr. 18.362)

René Werner*
(rene.werner@web.de)

Wintersemester 2003/2004

* Studiengang Informatik - 9.Fachsemester - Matr.nr. 5212832

Zusammenfassung

Gegenstand dieser Seminararbeit ist die Darstellung von logik-basierten Formalisierungen des Wissens- bzw. Glaubensbegriffs (epistemische Logiken), die in Multiagentensystemen etabliert sind. Dieses sind einerseits modallogische Systeme, deren Semantik auf dem von Hintikka geprägten Begriff der möglichen Welten basiert, und andererseits Systeme, die auf syntaktischer Ebene von Metasprachen Gebrauch machen (auf semantischer Ebene existieren unterschiedliche Ansätze).
Die Definition von Syntax und Semantik eines dem ersten Ansatz entsprechenden logischen Systems, eine Interpertation dieses im Kontext von Multiagentensystemen sowie eine Diskussion diesbzgl. bilden den Schwerpunkt der Arbeit. Es werden unter anderem gebräuchliche Axiomatierungen sowie das Auftreten des Problems der logischen Allwissenheit diskutiert. Motivation und Umsetzung des zweiten Ansatzes werden lediglich kurz skizziert.

Diese Arbeit folgt in weiten Teilen der Doktorarbeit von Michael J. Wooldridge:

> *The Logical Modelling of Computational Multi-Agent Systems*
> http://www.csc.liv.ac.uk/∼mjw/pubs/thesis.pdf (23.03.2004)

Inhaltsverzeichnis

1 Einleitung

Inhalt:

> „Agentenorientierung gehört zu den bedeutendsten aktuellen Forschungsströmen der Informatik. Eine besondere Herausforderung stellt dabei die Koordination autonomer Agenten dar. Wesentliche Grundlagen stammen aus dem Bereich der formalen Logik - z.B. Modal- und Temporallogiken, die das Weltwissen beschreiben bzw. die Handlungspläne erstellen. (...)“

Obiges ist in dem Kommentierten Veranstaltungsverzeichnis des Fachbereichs Informatik für das Wintersemester 2003/2004 (auszugsweise) als Inhalt der Lehrveranstaltung „Logiken für Multiagentensysteme“ angegeben und diente der Motivation zu dieser Seminararbeit.
Der Schwerpunkt der Arbeit liegt auf der logik-basierten Formalisierung des Wissensbegriffs zur Beschreibung des (Welt-)Wissens im Kontext eines Multiagentensystems. Eine Definition des Agentenbegriffs bzw. des Begriffs des Multiagentensystems wird nicht angegeben, hierzu sei auf die einleitenden Vorträge der genannten Lehrveranstaltung und zugehörige Ausarbeitungen verwiesen[1].

Die Arbeit gliedert sich wie folgt:
Anschließend an eine Diskussion des Wissensbegriffs erfolgt eine zunächst einführende Darstellung grundsätzlicher Repräsentationsansätze (Kap. 1.1 - 1.3). In dem Kap. 2 wird hierauf aufbauend ein modallogischer Ansatz vertiefend betrachtet, der im gegebenen Kontext etabliert ist. Abschließend erfolgt im Ausblick (Kap. 3) eine Beurteilung des zuvor definierten Systems im Hinblick auf eine umfassende Agentenmodellierung bzw. die Modellierung eines Multiagentensystems (die Modellierung des Wissensbegriffs ist nur ein Teil dieser; Planung, Strategien oder gar soziales Verhalten von Agenten sind weitere Aspekte).

[1] http://www.informatik.uni-hamburg.de/TGI/lehre/vl/WS0304/logik-fuer-mas/logik-fuer-mas.html

1.1 Wissen/Glauben

Der Begriff „epistemisch“ entstammt dem Griechischen: *episteme* = Wissen, Geschicklichkeit, Wissenschaft. Der für diese Seminararbeit gewählte Titel „Epistemische Logiken“ soll verdeutlichen, dass der Schwerpunkt der Arbeit auf der Darstellung von Logiken liegt, deren Zielsetzung die Formalisierung des Wissensbegriffs ist. Allerdings läßt sich der Begriff im Sinne einer eindeutigen Charakterisierung / Definition nur schwer fassen:

> „In order to formally reason about knowledge, we need a good semantic model. Part of the difficulty in providing such a model is that there is no agreement on exactly what the properties of knowledge are or should be. For example, is it the case, that you know what facts you know? Do you know what you don't know? Do you know only true things (...) ?“
>
> [J.Y. Halpern]

Eine Beantwortung der aufgeführten (und gerade in philosophischen Kreisen oftmals diskutierten) Fragen käme einer Definition des Wissensbegriffs nahe. Jedoch erscheint (mir) in diesem Zusammenhang folgendes Zitat zutreffend:

> „Für jede bekannte Definition [des Begriffs „Wissen“] gibt es Fälle, in denen sie offensichtlich nicht das wiedergibt, was wir unter Wissen verstehen.“
>
> [Wikipedia]

In Kap. 2.2.2 wird diese Diskussion aufgegriffen, wobei sich zeigen wird[2], dass sich im gegebenen Kontext die Ansicht etabliert hat, dass im Sinne der Modellierung des Wissensbegriffs sämtliche obige Fragen positiv zu beantworten seien. Wird hingegen die letzte Frage negativ beantwortet, so wird von „Glauben“ gesprochen, d.h. die Abgrenzung des Wissensbegriffs von dem Begriff des Glaubens kann wie folgt beschrieben werden:

> „Wissen steht seit der griechischen Philosophie im Gegensatz zur Meinung [Glauben]. Demnach impliziert Wissen Wahrheit und kann durch keine Argumentation widerlegt werden, während eine Meinung [Glauben] zwar wahr sein kann, aber diskutabel ist.“
>
> [Wikipedia]

1.2 Repräsentationsansätze

Die Entwicklung eines logischen Systems besteht aus zwei wesentlichen Teilen: der Spezifikation einer dem System zugrundeliegenden formalen Sprache (Syntax) und der Spezifikation einer zugehörigen Semantik, d.h. der Festlegung der Bedeutung der syntaktisch korrekten Ausdrücke. Im Folgenden wird beschrieben, welche Ansätze für diese beiden Ebenen im Rahmen einer formal-logischen Umsetzung des Wissensbegriffs Anwendung finden.

[2] [Wooldridge (1992)] folgend

Syntax: Auf syntaktischer Ebene haben sich zwei Ansätze etabliert:

1. Verwendung von Modaloperatoren:
   Die einem (etablierten) logischen System zugrundeliegende formale Sprache wird um Operatoren erweitert, die (angewendet auf logische Formeln) nicht wahrheitsfunktional wirken (siehe Kap. 2.1.1).

2. Verwendung einer Metasprache:
   Möchte man den Satz „Janine believes Cronos is the father of Zeus" mittels einer (prädikaten-)logischen Formel repräsentieren, liegt folgender Versuch einer Übersetzung nahe[3]: `Bel(Janine,Father(Zeus,Cronos))`.
   Nun ist aber `Father` ein Prädikatensymbol, demzufolge ist `Father(Zeus,Cronos)` eine prädikatenlogische Formel (und insbesondere kein Term) und somit im Sinne eines wohldefinierten Ausdrucks nicht als Argument des Prädikatensymbols `Bel` zu verwenden. Ein Ausweg ist die Einführung einer Metasprache, die Terme enthält, welche Formeln einer Objekt-Sprache denotieren (z.B. könnte die Metasprache einen Term ⌜`Father(Zeus,Cronos)`⌝ enthalten, welcher die objektsprachliche Formel `Father(Zeus,Cronos)` denotiert).
   Der Ansatz der Verwendung von Metasprachen wird im Rahmen dieser Arbeit nicht weiter verfolgt werden. Es sei allerdings darauf hingewiesen, dass als grundlegender Vorteil dieses Ansatzes häufig die Ausdrucksmächtigkeit genannt wird (z.B. ist es nicht möglich, die metasprachliche Formeln für „(Agent) i believes something" oder „(Agent) i believes everything (agent) j believes" mittels (quantifizierter) Modallogik auszudrücken. Allerdings birgt dieser Ansatz auch eine Reihe von Problemen, bzgl. deren Diskussion auf [Wooldridge (1992)] verwiesen sei.

Semantik: Auf der semantischen Ebene haben sich ebenfalls zwei Ansätze etabliert:

1. Possible Worlds Sematics (Semantik der Möglichen Welten):
   Dieser Ansatz basiert auf der Idee, das Wissen/Glauben (z.B. eines Agenten) als Menge möglicher Welten zu repräsentieren, die mittels einer Erreichbarkeitsrelation verknüpft sind (siehe hierzu Kapitel 1.3).

2. Interpreted symbolic Structures Approach:
   Das Wissen ist mittels symbolischer Formeln explizit in einer Datenstruktur repräsentiert, die mit dem Agenten assoziiert ist. Ein Agent weiß Aussage ϕ, falls ϕ in der mit dem Agenten assoziierten Datenstruktur enthalten ist. Für eine Diskussion dieses Ansatzes sei auf entsprechende Literatur verwiesen[4].

1.3 Mögliche Welten

Der verbreitetste und im Rahmen dieser Arbeit verfolgte Ansatz zur Modellierung der Annahmen eines Agenten über die Welt (d.h. seines Weltwissens) ist die Verwendung der

[3] nach **Genesereth, M., Nilsson, N.**, 1987: *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann Publishers.

[4] z.B. **Konolige, K.**, 1986: *A Deduction Model of Belief*. Pitman Publishing: London and Morgan Kaufmann.

Semantik der Möglichen Welten, meist verknüpft mit der Einführung von Modaloperatoren auf der syntaktischen Ebene. Dieser Ansatz zeichnet sich durch die Unabhängigkeit von der kognitiven Struktur des Agenten sowie einer ausgereiften mathematischen Theorie (bzgl. der Formalisierung) aus.
Der Begriff der möglichen Welten (engl.: possible worlds) wurde ursprünglich von Hintikka[5] geprägt. Zur Veranschaulichung diene folgendes Szenario:
Man betrachte einen skatspielenden Agenten. Zweifellos ist in diesem Spiel das Wissen über die Verteilung der Karten von unschätzbarem Wert, doch ist ein vollständiges Wissen diesbzgl. (bei ehrlicher Spielweise) ausgeschlossen. Folglich ist die Qualität eines Skatspielers (Agenten) in erster Linie (neben dem notwendigen Quäntchen Glück) durch die Fähigkeit bestimmt, aus gegebenen (und zu Beginn eines Spiels notwendigerweise begrenzten) Informationen Schlußfolgerungen über die Verteilung der Karten anzustellen. Hierzu kann der Agent nun alle möglichen Kombinationen der Kartenverteilung berechnen und sich diese notieren. Einige dieser Kombinationen sind allerdings unzutreffend, da sie nicht mit dem Wissen des Agenten übereinstimmen (z.B. sind sämtliche Kombinationen auszuschließen, in denen Karten, die der Agent in seiner Kartenhand hält, anderen Spielern zugeteilt sind). Alle anderen Kombinationen sind epistemische Alternativen (mögliche Welten), d.h. Kombinationen, die basierend auf dem aktuellen Wissen des Agenten möglich sind. Der Agent weiß/glaubt genau das, was in allen möglichen Welten wahr ist (sollte zum Beispiel in allen möglichen Welten ein Spieler sämtliche Trümpfe halten, so wird der Agent dieses als wahr annehmen, d.h. er wird dieses wissen/glauben).

[5] in **Hintikka, J.**,1962: *Knowledge and Belief.* Cornell University Press.

2 Modallogik

Inhalt:

Modallogiken sind Erweiterungen der klassischen Logiken[1] um die Konzepte „Zustände“ (Welten) und „Zustandsübergänge“ (Erreichbarkeitsrelation, Erreichbarkeit von Welten)[2]: Die Operatoren der klassischen Logiken wirken wahrheitsfunktional, d.h. der Wahrheitswert einer mit ihrer Hilfe gebildeten Aussage hängt ausschließlich von den Wahrheitswerten der Argumente der Operatoren ab, welche bei gegebener Belegung fix sind. Mit Einführung von nicht-wahrheitsfunktionalen Modal-Operatoren wird eine Kontextabhängigkeit des Wahrheitswertes einer Aussage (d.h. die Abhängigkeit dieses von den Zuständen) zugelassen.
Die Interpretation einer Modallogik ist meist von der intendierten Anwendung abhängig und die Bezeichnungen der erweiternden Konzepte entsprechend unterschiedlich. Der Begriff des Zustandes ist nahezu allgemein anwendbar, während der Welten-Begriff eine Interpretation im Sinne einer Semantik der möglichen Welten nahelegt.

In Kap. 2.1 erfolgt die formale Darstellung einer entsprechend erweiterten Aussagenlogik, welche in Kap. 2.2 gemäß der intendierten Anwendung als „Logik des Glaubens/Wissens“ in einem Multiagentensystems (MAS) interpretiert, erweitert und diskutiert wird.

2.1 Modale Aussagenlogik

Die Definition eines logischen Systems erfolgt durch Spezifikation einer dem System zugrundeliegenden formalen Sprache (Syntax) und Spezifikation von Evaluations- bzw. Interpretationsprinzipien (Semantik), mittels welcher den Ausdrücken der formalen Sprache eine Bedeutung zugeordnet wird (z.B. in der Prädikatenlogik: Formeln $\rightarrow$ Wahrheitswerte, Terme $\rightarrow$ Individuen).
Hierüber hinaus erfolgt i.A. eine Spezifikation semantischer Kategorien sowie die Spezifikation von Beweisverfahren.

[1] Aussagenlogik, Prädikatenlogik erster Stufe
[2] vgl. [Nonnengart, Ohlbach (1992)]

Obigem entsprechend werden in diesem Kapitel Syntax und Semantik des Basissystems der modalen Aussagenlogik[3] (K-System, s. Kap. 2.1.3) definiert. Zugehörige Beweisverfahren werden nicht behandelt, bei Interesse sei z.B. auf [Hughes (1978)] verwiesen. Anschließend erfolgt eine (kurze) Einführung in ausgewählte, auf dem K-System basierende modallogische Systeme.

2.1.1 Syntax

Ausgehend von der klassischen, d.h. zweiwertigen Aussagenlogik erfolgt eine Erweiterung dieser um die Modaloperatoren $\Box$ und $\Diamond$, deren intendierte Bedeutung das Ausdrücken von „Notwendigkeit“ und „Möglichkeit“ der Wahrheit einer Aussage ist (s. Kap. 2.1.2).

Das Vokabular der modalen Aussagenlogik besteht aus einer abzählbaren Menge atomarer Formeln $At = \{p, q, r, \dots\} \cup \{\top, \bot\}$, wobei $\top$ und $\bot$ logische Konstanten und $\{p, q, r, \dots\} = As$ eine abzählbare Menge der atomaren Aussagen sind, sowie den aus der Aussagenlogik bekannten Junktoren $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$ und den Modaloperatoren $\Box$, $\Diamond$. Hieraus werden die modallogischen Formeln gebildet (vgl. [Schneider (2002)]):

Definition 2.1 (modallogische Formel) :

(1) Das Wort ϕ ist eine (modallogische) Formel, falls es aus den genannten Zeichen besteht und eine der folgenden Gestalten hat:

$\top$	
p	für ein bel. $p \in As$
$\neg\psi$	für eine bel. Formel ψ
$\psi_1 \wedge \psi_2$	für bel. Formeln ψ_1, ψ_2
$\Box\psi$	für eine bel. Formel ψ

(2) Weiterhin werden folgende Kurzschreibweisen vereinbart:

$$\phi \vee \psi = \neg(\neg\phi \wedge \neg\psi)$$
$$\phi \Rightarrow \psi = \neg\phi \vee \psi$$
$$\phi \Leftrightarrow \psi = (\phi \Rightarrow \psi) \wedge (\psi \Rightarrow \phi)$$
$$\bot = \neg\top$$
$$\Diamond\phi = \neg\Box\neg\phi$$

Die Modaloperatoren sind folglich dual zueinander.

2.1.2 Semantik

Die Semantik der klassischen Modallogik geht auf S. Kripke resp. der Leibnizschen Vorstellung des „Notwendigen“ zurück, welcher zufolge etwas notwendigerweise wahr ist, wenn es in allen vorstellbaren Welten wahr ist.Die Grundstruktur der Semantik ist das

[3] im Folgenden auch mit (klassischer) Modallogik bezeichnet, obgleich die Verwendung (je nach zugrundeliegender Logik) mächtigerer Modallogiken durchaus denkbar ist (s. Kap. 2.2.2, Abschnitt „Quantifizierte Modallogik“).

modallogische Modell, zu Ehren S. Kripkes auch Kripke-Struktur genannt, bestehend aus einem modallogischen (Kripke-) Rahmen und einer Belegungsfunktion:

Definition 2.2 (modallogisches Modell) :

Sei W eine nicht leere Menge von Welten und $R \subseteq W \times W$ eine binäre Relation auf W (Zugänglichkeits-, Sichtbarkeitsrelation). Dann:

(1) Als einen (modallogischer) Rahmen bezeichnet man folgende Struktur:

$F = \langle W, R \rangle$

(2) Eine Belegungsfunktion ordnet jeder Welt die Menge der in dieser wahren Aussagen zu:

$\pi : W \rightarrow 2^{At}$

(3) Als ein (auf einem Rahmen $F = \langle W, R \rangle$ basierendes) (modallogisches) Modell wird folgende Struktur bezeichnet:

$$\begin{aligned} M &= \langle F, R \rangle \\ &= \langle W, R, \pi \rangle \end{aligned}$$

Die eigentliche Semantik der Sprache im Sinne der Spezifikation eines Evaluationsprinzips, d.h. der Abbildung von (modallogischen) Formeln auf die Wahrheitswerte $true$ oder $false$, wird mittels einer Erfüllbarkeitsrelation $\models$ zwischen einem Tupel $\langle M, w \rangle$ und einer modallogischen Formel ϕ definiert.
Hierbei ist $w \in W$ als Referenzwelt zu betrachten, bzgl. welcher die Wahrheit von ϕ im Modell M mit Hilfe der folgenden semantischen Regeln gegeben ist (bzw. bestimmt werden kann). Hierbei ist eine Formel ϕ wahr (bzgl. eines Modells M) in w, falls gilt: $\langle M, w \rangle \models \phi$.

Definition 2.3 (Erfüllbarkeitsrelation) :

Sei $M = \langle W, R, \pi \rangle$ ein modallogisches Modell, $w \in W$. Dann wird eine (die Semantik der modalen Aussagenlogik festlegende) Erfüllbarkeitsrelation wie folgt definiert:

$$\begin{array}{lll} \langle M, w \rangle \models true & & \\ \langle M, w \rangle \models p & :\Leftrightarrow & p \in \pi(w) \\ \langle M, w \rangle \models \neg\phi & :\Leftrightarrow & \langle M, w \rangle \not\models \phi \\ \langle M, w \rangle \models \phi \vee \psi & :\Leftrightarrow & \langle M, w \rangle \models \phi \ \text{ oder } \ \langle M, w \rangle \models \psi \\ \langle M, w \rangle \models \Box\phi & :\Leftrightarrow & \forall w' \in W. \ \text{wenn } (w, w') \in R, \ \text{dann } \langle M.w' \rangle \models \phi \end{array}$$

wobei $p \in As$ und ϕ, ψ modallogische Formeln seien.

Neben dem Begriff der Wahrheit einer Formel in einer Welt eines Modells werden auch die Begriffe der Erfüllbarkeit sowie der Wahrheit in Modellen bzw. Rahmen oder Wahrheit als solcher verwendet:

Definition 2.4 :

Seien $F = \langle W, R \rangle$ ein modallogischer Rahmen, $M = \langle W, R, \pi \rangle$ ein auf $F = \langle W, R \rangle$ basierendes Modell und ϕ eine modallogische Formel. Dann:

(1) ϕ ist **erfüllbar in** M gdw. es eine Welt w von M gibt, so dass ϕ in w wahr ist.

(2) ϕ ist **erfüllbar** gdw. es ein Modell gibt, in welchem ϕ erfüllbar ist. w wahr ist.

(3) ϕ ist **wahr in** M gdw. ϕ in allen Welten $w \in W$ wahr ist.
$(M \models \phi)$

(4) ϕ ist **wahr in** F gdw. ϕ in allen auf F basierenden Modellen wahr ist.
$(F \models \phi)$

(5) ϕ ist **wahr(gültig)** gdw. ϕ in allen Rahmen wahr ist.
$(\models \phi)$

Der Begriff der Wahrheit einer Formel in einem Rahmen bedingt die Unabhängigkeit der Wahrheit dieser Formel von der Evaluationsfunktion der auf diesem Rahmen basierenden Modelle. Die Wahrheit einer solchen Formel kann allerdings von der Beschaffenheit der Erreichbarkeitsrelation (des Rahmens) abhängig sein. Insbesondere läßt sich die Gültigkeit von (im gegebenen Kontext relevanten, s. Kap. 2.1.3) Axiomen durch bestimmte Forderungen an die Erreichbarkeitsrelation erzielen[4].

[4]dieser Zusammenhang wird im Rahmen der Korrespondenztheorie behandelt, siehe z.B. **Blackburn, P., de Rijke, M., Venema, Y.**, 2001: *Modal Logic*. Cambridge Tracts in Theoretical Computer Science 53. Cambridge University Press.

2.1.3 Axiomatisierungen

Die bisher dargestellte Modallogik läßt sich durch die Gültigkeit des Axioms[5]

$$\Box(\phi \Rightarrow \psi) \Rightarrow (\Box\phi \Rightarrow \Box\psi)$$

sowie die beiden Inferenzregeln

$$\text{wenn} \models (\phi \Rightarrow \psi) \text{ und} \models \phi, \text{ dann} \models \psi$$
$$\text{wenn} \models \phi, \text{ dann} \models \Box\phi$$

charakterisieren.
Die erste Inferenzregel stellt den Modus Ponens dar, die zweite Regel wird als Notwendigkeitsregel bezeichnet. Das obige Axiom wird (wiederum nach S. Kripke) K-Axiom genannt. Gemäß der Betrachtung eines logischen Systems als eine Menge von Axiomen (deren Elemente als Theoreme bezeichnet werden) und Inferenzregeln spricht man in diesem Zusammenhang von einem K-System. Da das K-Axiom ein Theorem jeder vollständigen Axiomatisierung modaler Aussagenlogik ist, bildet das K-System das Basissystem dieser.

Im Rahmen spezieller Anwendungen (z.B. Modallogik als epistemische Logik (im MAS)) können zusätzliche Anforderungen an die Eigenschaften des modallogischen Systems gestellt sein. Diesbzgl. werden, basierend auf dem K-System, größere Systeme gebildet, indem im K-System nicht beweisbare Formeln dem System als Axiome hinzugefügt werden.
Die im gegebenen Kontext wichtigsten Axiome mit der jeweils korrespondierenden Eigenschaft sind im Folgenden aufgeführt:

K-Axiom: $\Box(\phi \Rightarrow \psi) \Rightarrow (\Box\phi \Rightarrow \Box\psi)$
(R beliebig)

T-Axiom: $\Box\phi \Rightarrow \phi$
(R reflexiv, d.h.:
$\forall w \in W.(w, w) \in R$)

D-Axiom: $\Box\phi \Rightarrow \Diamond\phi$
(R seriell, d.h.:
$\forall w \in W.\exists w' \in W.(w, w') \in R$)

4-Axiom: $\Box\phi \Rightarrow \Box\Box\phi$
(R transitiv, d.h.:
$\forall w, w', w'' \in R.(w, w') \in R \wedge (w', w'') \in R \Rightarrow (w, w'') \in R$)

5-Axiom[6]: $\Diamond\phi \Rightarrow \Box\Diamond\phi$
(R euklidisch, d.h.:
$\forall w, w', w'' \in R.(w, w') \in R \wedge (w, w'') \in R \Rightarrow (w', w'') \in R$)

[5] genau genommen handelt es sich um ein Axiomenschema, da ϕ, ψ beliebige (modallogische) Formeln sind

[6] in der Literatur auch als E-Axiom zu finden

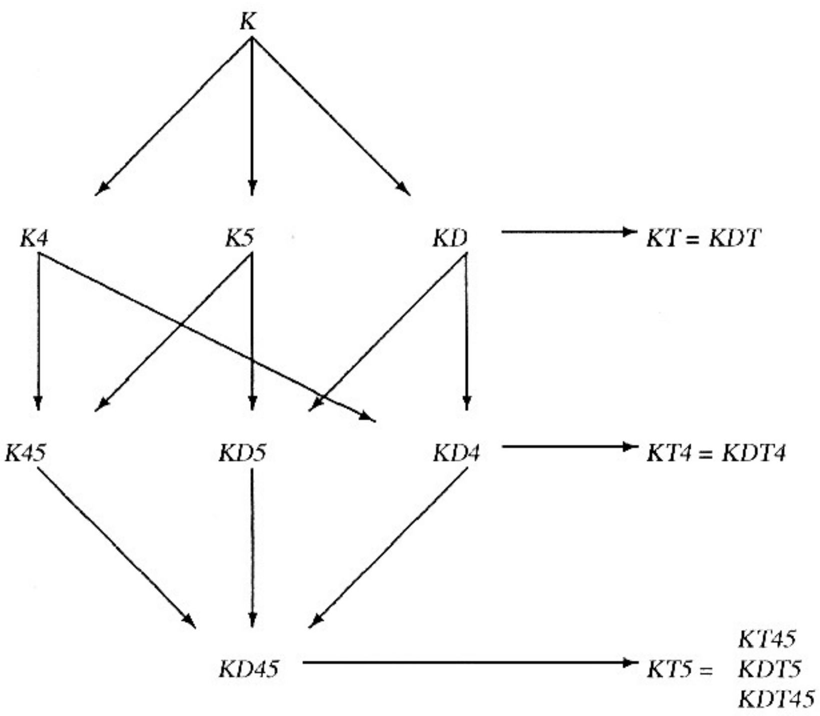


Abbildung 2.1: Hierarchie der auf den Axiomen K, D, T, 4 und 5 basierenden modalen Systemen (Quelle: [Wooldridge (1992)])

Für Beweise dieser und weiterer Korrespondenzsysteme sei auf entsprechende Literatur verwiesen[7].

Je nach Axiomatisierung werden modale Systeme mit $K\Sigma_1 \ldots \Sigma_n$ bezeichnet, wobei $\Sigma_1, \ldots, \Sigma_n$ die Bezeichnungen der das System charakterisierenden Axiome sind. Für einige Systeme sind darüber hinaus Abkürzungen festgelegt worden:

$$\begin{array}{llll} \text{T} & := \text{KT} & = & \text{K} + \text{T} \\ \text{S4} & := \text{KT4} & = & \text{K} + \text{T} + 4 \\ \text{weak-S5} & := \text{KD45} & = & \text{K} + \text{D} + 4 + 5 \\ \text{S5} & := \text{KT5} & = & \text{K} + \text{T} + 5 \end{array}$$

Die Auswahl abzukürzender Systemnamen ist insbesondere in der häufigen Verwendung dieser Systeme begründet. Zudem erweist sich, dass aus den Axiomen T, D, 4 und 5 lediglich 11 (statt $2^4 = 16$) unterschiedliche[8] Systeme gebildet werden können und die definierten Abkürzungen T, S4, S5 je mehr als ein System bezeichnen. Die Hierarchie der Systeme ist Abb. 2.1 zu entnehmen. Ein Pfeil von System A zu System B bedeutet, dass System B eine strikte Obermenge von System A ist (jedes Theorem aus A gilt in B, die Umkehrung gilt nicht), A = B bedeutet, dass A und B äquivalente Systeme sind.

[7] z.B. Blackburn, P., de Rijke, M., Venema, Y., 2001: *Modal Logic*. Cambridge Tracts in Theoretical Computer Science 53. Cambridge University Press.

[8] ein System unterscheidet sich von einem anderen System, falls die zugehörigen Theoreme nicht übereinstimmen

2.2 Modallogik als Epistemische Logik in einem MAS

Inhalt dieses Kapitels ist die Interpretation der Modallogik im Sinne einer epistemischen Logik in einem Multiagentensystem (hierzu erfolgt eine Erweiterung der bisher dargestellten Logik, s. Kap. 2.2.1) mit anschließender Diskussion dieser (Kap. 2.2.2).

2.2.1 Epistemische Interpretation

In einer epistemischen Interpretation der Modallogik verwendet man den $\Box$-Operator[9], um den Glauben (die Überzeugung) resp. das Wissen eines Agenten darzustellen. $\Box\phi$ wird dann gelesen als „der (betrachtete) Agent glaubt/weiß ϕ“. Die Welten des Modells werden als epistemische Alternativen interpretiert, wobei die von einer Welt aus erreichbaren Alternativen über die Zugänglichkeitsrelation festgelegt werden. Wissen bzgl. ϕ liegt vor, falls ϕ in allen von dem aktuellen Zustand / der aktuellen Welt aus erreichbaren Alternativen wahr ist.
Dieser Ansatz läßt sich für ein Multiagentensystem[10] wie folgt erweitern:

Syntaxerweiterung: $\Box \rightsquigarrow \{\mathrm{K}_1, \ldots, \mathrm{K}_n\}$
Die formale Sprache wird durch die Ersetzung des $\Box$-Operators durch eine Menge $\{\mathrm{K}_1, \ldots, \mathrm{K}_n\}$ von einstelligen Wissens-Operatoren erweitert. Die Formel $\mathrm{K}_i\phi$ wird als „(Agent) i glaubt/weiß ϕ“ gelesen.

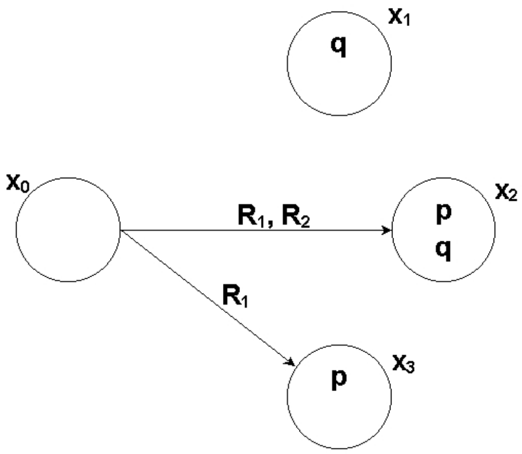
Anpassung der Semantik: $\langle W, R, \pi\rangle \rightsquigarrow \langle W, R_1, \ldots, R_n, \pi\rangle$
Im oben betrachteten Fall eines einzelnen Agenten ist das Wissen in einem bestimmten Zustand mittels der Erreichbarkeitsrelation ableitbar, d.h. das Wissen eines Agenten ist im Wesentlichen durch diese gegeben. In Analogie hierzu wird für ein Multiagentensystem für jeden Agenten eine Erreichbarkeitsrelation definiert. Das zugehörige Modell hat dann folgende Struktur: $M = \langle W, R_1, \ldots, R_n, \pi\rangle$
Hierauf basierend wird die semantische Regel für den $\Box$-Operator durch folgende Regel ersetzt (alle anderen Regeln bleiben unverändert):

$$\langle M, w\rangle \models \mathrm{K}_i\phi \quad :\Leftrightarrow \quad \forall w' \in W.\, \text{wenn}\, (w, w') \in R_i,\, \text{dann}\, \langle M, w'\rangle \models \phi$$

Darstellung allgemeinen, gemeinsamen und verteilten Wissens: EK, CK, DK
Im Kontext von Multiagentensystemen ist es zudem oftmals erstrebenswert, Aussagen über das Wissen der Gesamtheit der Agenten bzw. das Gesamtwissen der Agenten treffen zu können.
Diesbzgl. unterscheidet man i.A. folgende drei Arten des Wissens:

1.) allgemeines Wissen
2.) gemeinsames Wissen
3.) verteiltes Wissen

[9] der $\Diamond$-Operator, ohnehin nur als abkürzende Schreibweise eingeführt, wird nicht weiter betrachtet
[10] hier und im Weiteren wird von einer endlichen Anzahl von Agenten ausgegangen



Nebenstehendes Modell (für ein Zwei-Agenten-System) enthält die Welten $x_0, \ldots, x_3$. Die beschrifteten Kanten visualisieren die Erreichbarkeitsrelationen. Betrachtet man x_0 als Referenzwelt, so weiß Agent 2, dass die Aussagen p und q wahr sind, da beide Aussagen in den für Agent 2 erreichbaren Welten (hier nur Welt x_2) wahr sind. Agent 1 hingegen kann sich ob der Wahrheit von Aussage q nicht sicher sein, da diese in der für ihn erreichbaren Welt x_3 nicht wahr ist.

Abbildung 2.2: Beispiel zur Veranschaulichung des Modellbegriffs

---

Allgemeines Wissen sind die Aussagen, von denen jeder Agent weiß, dass sie wahr sind (EK $\hat{=}$ everyone knows):

$$\mathrm{EK}\phi \;:=\; \mathrm{K}_1\phi \wedge \cdots \wedge \mathrm{K}_n\phi$$

Die entsprechende semantische Regel lautet:

$$\langle M, w\rangle \models \mathrm{EK}\phi \;:\Leftrightarrow\; \forall i \in \{1, \ldots, n\}. \langle M, w\rangle \models \mathrm{K}_i\phi$$

Gemeinsames Wissen sind die Aussagen, von denen jeder weiß, dass sie wahr sind, und zudem jeder weiß, dass jeder weiß, dass sie wahr sind, und so fort. Allgemeines und gemeinsames Wissen beschreiben folglich nicht das selbe, insbesondere ist gemeinsames Wissen nicht immer erreichbar[11]. Der zugehörige Operator wird wie folgt definiert:

$$\mathrm{CK}\phi \;:=\; \mathrm{EK}^1\phi \wedge \mathrm{EK}^2\phi \wedge \cdots \wedge \mathrm{K}^k\phi \wedge \ldots$$

wobei EK^k zu lesen ist als „jeder (Agent) glaubt/weiß ϕ bis zum Grad k" und rekursiv definiert ist (CK $\hat{=}$ common knowledge):

$$\begin{aligned} \mathrm{EK}^1\phi &:= \mathrm{EK}\phi \\ \mathrm{EK}^{k+1}\phi &:= \mathrm{EK}(\mathrm{EK}^k\phi) \end{aligned}$$

Die zugehörige semantische Regel lautet:

$$\langle M, w\rangle \models \mathrm{CK}\phi \;:\Leftrightarrow\; \forall k \in \mathbb{N}^*. \langle M, w\rangle \models \mathrm{EK}^k\phi$$

Als verteiltes Wissen (implizites Wissen) wird das Gesamtwissen in dem Agentensystem bezeichnet. Es ist das Wissen, welches ein allwissender Beobachter des Agentensystems haben würde. Der zugehörige Operator kann nicht aus den bereits definierten Operatoren

[11] eine Veranschaulichung dieses Problems bietet das „coordinated attack problem", siehe z.B. **Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.**, 1994: *Reasoning About Knowledge*. MIT Press.

abgeleitet werden (vgl. hierzu [WOOLDRIDGE (1992)], s. 20). Seine Definition erfolgt über folgende semantische Regel (DK $\hat{=}$ distributed knowledge):

$$\langle M, w \rangle \models \mathrm{DK}\phi \quad :\Leftrightarrow \quad \forall w' \,\mathrm{mit}\, (w, w') \in (R_1 \cap \cdots \cap R_n).\, \langle M, w' \rangle \models \phi$$

Die definierten Wissensoperatoren bilden folgende Hierarchie:

$$\mathrm{CK}\phi \Rightarrow \cdots \Rightarrow \mathrm{EK}^k\phi \Rightarrow \cdots \Rightarrow \mathrm{EK}^2\phi \Rightarrow \mathrm{EK}\phi \Rightarrow \mathrm{K}_i\phi \Rightarrow \mathrm{DK}\phi$$

Für eine weitergehende Diskussion dieser Operatoren sei auf [HALPERN, MOSES (1996)] verwiesen.

2.2.2 Diskussion

Nach erfolgter Definition einer in Syntax und Semantik dem Anwendungsfeld Multiagentensystem angepassten Modallogik wird in diesem Kapitel die Frage nach einer geeigneten Axiomatisierung zu verwendender Systeme und der Angemessenheit der Anwendung eines modallogischen Systems als epistemische Logik diskutiert.
Der Schwerpunkt liegt auf der Darstellung des „Problems der logischen Allwissenheit“, welches i.A. als ein wesentlicher Kritikpunkt bzgl. der Verwendung eines modallogischen Ansatzes zur Formalisierung des Wissensbegriffs angesehen wird.
Weiterhin erfolgt eine Betrachtung der Erweiterung der bislang dargestellten Modallogik um das Konzept der Quantifizierung.

Axiomatisierung:

> „[Is] it the case, that you know what facts you know? Do you know what you don’t know? Do you know only true things, or can something you „know“ actually be false?“ [J.Y. HALPERN]

Bei der Beantwortung der Frage nach einer passenden Axiomatisierung einer „Logik des Glaubens/Wissens“ erfolgt in der Regel eine Beschränkung auf die Betrachtung der in Kap. 2.1.3 eingeführten Axiome[12]. Diese können im gegebenen Kontext wie folgt interpretiert werden (das K_n-Axiom wird im nächsten Abschnitt behandelt):

D_n-Axiom[13]: $\quad \mathrm{K}_i\phi \Rightarrow \neg\mathrm{K}_i\neg\phi$

Widerspruchsfreiheit: „Falls (Agent) i ϕ glaubt/weiß, dann glaubt/weiß er nicht $\neg\phi$.“

T_n-Axiom: $\quad \mathrm{K}_i\phi \Rightarrow \phi$

Wissensaxiom: „Die Aussagen, die (Agent) i glaubt/weiß, sind wahr.“

4_n-Axiom: $\quad \mathrm{K}_i\phi \Rightarrow \mathrm{K}_i\mathrm{K}_i\phi$

Positive Introspektion: „(Agent) i glaubt/weiß, was er glaubt/weiß.“ (der Agent ist sich

[12] vgl. z.B. [SCHNEIDER (2002)]

[13] der Notation von [WOOLDRIDGE (1992)] folgend werden die ein n-Agenten-MAS charakterisierenden Axiome mit Σ_n bezeichnet

seiner Annahmen bewußt)

5_n-Axiom: $\neg K_i \neg \phi \Rightarrow \Box \neg K_i \neg \phi$

Negative Introspektion: „(Agent) i glaubt/weiß, was er nicht glaubt/weiß." (im Vergleich zum 4-Axiom ist die Gültigkeit dieses Axioms eine stärkere Bedingung an das System)

Prinzipiell beschränkt die Gültigkeit von Axiomen die Menge der erreichbaren epistemischen Alternativen, wodurch die Unsicherheit bzgl. des Wissens/Glaubens des Agenten reduziert wird[14]. Betrachtet man dieses als erstrebenswert, reduziert sich die Frage nach einer angemessenen Axiomatisierung auf die Frage, ob das zu definierende System eher den Begriff des Glaubens oder den Begriff des Wissens modellieren soll (siehe hierzu Kap. 1.1). Dieses entspricht gemäß obigen Ausführungen der Frage nach der Gültigkeit des T-Axioms. Entsprechend haben sich folgende Systeme etabliert: $S5_n$ ($= K_n T_n 5_n = K_n D_n T_n 4_n 5_n$) als Logik des (idealisierten) Wissens, weak-$S5_n$ ($=K_n D_n 4_n 5_n$) als Logik des (idealisierten) Glaubens.

Angemessenheit, Problem der logischen Allwissenheit: Zur Beurteilung der Angemessenheit der dargestellten Modallogik als Logik des Wissens und/oder des Glaubens ist zunächst das K-Axiom (bzw. im MAS entsprechend das K_n-Axiom) und dessen Zusammenwirken mit der Notwendigkeitsregel zu betrachten, da beide in jedem modallogischen System enthalten sind (s. Kap 2.1.3):

K_n-Axiom: $K_i(\phi \Rightarrow \psi) \Rightarrow (K_i \phi \Rightarrow K_i \psi)$

Dieses besagt im Wesentlichen (unter Berücksichtigung des Modus Ponens), dass das Agentenwissen unter logischer Implikation abgeschlossen ist.

Notwendigkeitsregel: wenn $\models \phi$, dann $K_i \phi$

Im Kontext eines MAS bedeutet dieses, dass jeder Agent alle gültigen Formeln (insbesondere auch alle aussagenlogischen Tautologien) kennt. Gemeinsam folgt, dass jeder Agent alle (aussagenlogischen) Konsequenzen kennt, die sich aus seinem Wissen ergeben. Dieses bezeichnet man als das „Problem der logischen Allwissenheit" (engl.: Logical Omniscience Problem, LOP).
Insgesamt erscheint Obiges im Sinne der Agentenmodellierung unangemessen in Anbetracht der grundsätzlich endlichen Ressourcen (Speicherkapazitäten) bzgl. der Umsetzung dieser. Zudem wirkt der Abschluß des Wissens unter logischer Implikation kontraintuitiv; beispielsweise gelingt es in der Mathematik oftmals eine lange Zeit nicht, Sätze zu beweisen, die prinzipiell aus bekanntem Wissen folgerbar sind (Fermats letzter Satz, ...).
Des Weiteren ergibt sich als Konsequenz aus Obigem, dass das Arbeiten mit inkonsistentem Wissen nicht möglich ist, da Inkonsistenz zur Folge hätte, dass alle syntaktisch korrekten Ausdrücke als Wissen abgeleitet werden könnten (aus einer inkonsistenten Wissensbasis läßt sich alles folgern!). Dieses erscheint (zumindest aus Sicht eines menschlichen Agenten) wenig intuitiv.

[14] es gilt bzgl. der Theoreme (Tautologien) der entstehenden Systeme folgende Hierarchie:
$\mathrm{Taut}(K) \subsetneq \mathrm{Taut}(T) \subsetneq \mathrm{Taut}(S_4) \subsetneq \mathrm{Taut}(S_5)$

Levesque sah sich deshalb zu folgender Beurteilung der Angemessenheit eines wie vorhergehend definierten modallogischen Systems bzgl. der Formalisierung des Glaubens-/Wissens-Begriffs veranlaßt:

> „Any one of these [problems] might cause one to reject a possible-world formalization as unintuitive at best and completly unrealistic at worst“
>
> [Levesque]

Ansätze zur Lösung des Problems der logischen Allwissenheit (d.h. der Definition einer epistemischen Logik auf Basis der Semantik der möglichen Welten unter Vermeidung des Auftretens des LOP's) werden im Rahmen des Ausblicks (Kap. 3) skizziert.

Quantifizierte Modallogik: Bzgl. der Einführung von Modalitäten in ein bestehendes logisches System ist man nicht auf die klassische Aussagenlogik beschränkt (s. Seite 7, erster Abs.). Insbesondere im Bereich der Wissenrepräsentation scheint es nahezuliegen, auch Quantifizierung zu gestatten, d.h. als logische Grundsprache die klassische Prädikatenlogik zu wählen. Syntaktisch bedeutet dieses in Analogie zur Einführung von Modalitäten in die Aussagenlogik (s. Kap. 2.1) eine Erweiterung der der Prädikatenlogik zugrundeliegenden formalen Sprache (wird als bekannt vorausgesetzt, andernfalls sei z.B. auf [Habel, Eschenbach (2003)] verwiesen) um die Modaloperatoren $\Box$, $\Diamond$. Ausgehend von dem modallogischen Modell-Begriff (s. Kap. 2.1.2)

$$M = \langle W, R, \pi \rangle$$

und dem Begriff des prädikatenlogischen Modells

$$M = \langle D, I \rangle$$

wobei:

D ist eine bel. nicht leere Menge (Domäne)
I ist eine Abbildung (Interpretation) mit:

für jede Konstante c gilt: $I(c) \in D$
für jedes n-stellige Funktionssymbol f gilt: $I(f) : D^n \to D$
für jedes n-stellige Relationssymbol P gilt: $I(P) \subseteq D^n$

sowie dem Begriff der Variablenbelegung gilt es folglich, die beiden Sematikkonzeptionen zu kombinieren.
Ein möglicher Ansatz ist die weltspezifische Interpretation/Auswertung von Konstanten, Funktions- und Prädikatssymbolen, wobei die Interaktion von Quantifikation und Modalität durch den ggf. nötigen Zugriff auf unterschiedliche Domänen gegeben ist. Varianten bzgl. der Umsetzung sind z.B. die Wahl einer für alle Welten konstanten Domäne[15], eine Abhängigkeit der Domäne von den Welten, die Forderung nach Monotonie bzgl. der Zugänglichkeit (d.h. $R(w_1, w_2)$, dann $D(w_1) \subseteq D(w_2)$) und die Forderung von Rigidität (d.h. Konstanteninterpretation unabhängig von der Welt).

[15] dieses ist sicherlich die naheliegendste Wahl. Allerdings folgt die Gültigkeit der Barcan-Schemata, welche - insbesondere aufgrund der intendierten Verwendung der Logik - nicht von Vorteil ist (s. [Wooldridge (1992)], S. 21)

Die Wahl der Prädikatenlogik als Basis der „Logik des Glaubens/Wissens“ ist folglich mit erheblichem Mehraufwand bzgl. der Defintion einer entsprechenden Semantik verbunden. Darüberhinaus ist auch ein solches logisches System nicht unbedingt zur Formalisierung des Glaubens-/Wissensbegriffs geeignet. Man betrachte hierzu den folgenden Satz:

It is known that there is a unicorn.

Mögliche Repräsentationen sind:

$\exists x.\Box \text{Unicorn}(x)$

$\Box \exists x.\text{Unicorn}(x)$

Keine der beiden obigen Formeln scheint geeignet, die Aussage des obigen Satzes auszudrücken. Darüberhinaus unterscheiden sich beide Formeln bzgl. dessen, was ausgesagt wird (der Schwerpunkt der ersten Formel liegt auf der Existenz eines dem Agenten bekannten und von ihm zu benennden Individuums, welches die Eigenschaft, ein Einhorn zu sein, erfüllt; die zweite Formel hingegen besagt nicht, dass der Agent ein solches Individuum benennen kann).

3 Ausblick

Inhalt:

Ein zentraler Kritikpunkt bzgl. der logik-basierten Formalisierung des Wissens- bzw. Glaubensbegriffs mittels einer Semantik der möglichen Welten (basierend auf der Einführung von Modaloperatoren, s. Kap. 2) ist das in Kap. 2.2.2 beschriebene, aus dem Zusammenspiel von K-Axiom und Notwendigkeitsregel resultierende Problem der logischen Allwissenheit (LOP). In Kap. 3.1 werden zwei Ansätze skizziert, die - unter Beibehaltung der grundsätzlichen Struktur einer Semantik der möglichen Welten - versuchen, dieses Problem zu lösen. In Kap. 3.2 erfolgt eine abschließende Betrachtung des Dargestellten im Sinne einer Einordnung im Kontext der Agentenmodellierung.

3.1 LOP: Lösungsansätze

3.1.1 Levesque: Logic of implizit and explicit belief

Der von Levesque vorgeschlagene Lösungsansatz[1] basiert auf der Unterscheidung von explizitem und implizitem Wissen. Für beide Wissensarten wird ein entsprechender Operator eingeführt: B für explizites Wissen und L für implizites Wissen. Eine Verschachtelung dieser Operatoren ist nicht zulässig.
Zielsetzung ist die Definition einer Semantik, bzgl. derer explizites Wissen nicht unter logischer Schlussfolgerung abgeschlossen ist. Hierzu führt Levesque den Begriff der Situation ein. In einer Situation ist - im Gegensatz zu dem Begriff der Welt - eine Aussage nicht zwangsläufig wahr oder falsch, sie kann auch keines von beidem oder sogar beides sein (inkohärente Situation). Ein Modell dieser Logik hat folgende Gestalt:

$$M = \langle S, B, T, F \rangle$$

wobei:

S	Menge der Situationen
$B \subseteq S$	Menge derjenigen Situationen, die mit dem expliziten Wissen des Agenten übereinstimmen
$T : At \rightarrow 2^S$	Abbildung, die einer atomaren Aussage diejenigen Situationen zuordnet, in denen dieser der Wahrheitswert *true* zugeordnet ist

[1] siehe **Levesque, H.**, 1984: *A logic of implicit and explicit belief.* In: AAAI-84

$F : At \to 2^S$	Abbildung, die einer atomaren Aussage diejenigen Situationen zuordnet, in denen dieser der Wahrheitswert $false$ zugeordnet ist

Die Semantik selbst wird mittels zweier Erfüllbarkeitsrelationen und folgender semantischer Regeln festgelegt:

$$
\begin{array}{lll}
\langle M, s\rangle \models_T p & :\Leftrightarrow & s \in T(s) \text{ (wobei } p \in At) \\
\langle M, s\rangle \models_F p & :\Leftrightarrow & s \in F(s) \text{ (wobei } p \in At) \\
\langle M, s\rangle \models_T \phi \vee \psi & :\Leftrightarrow & \langle M, s\rangle \models_T \phi \text{ oder } \langle M, s\rangle \models_T \psi \\
\langle M, s\rangle \models_F \phi \vee \psi & :\Leftrightarrow & \langle M, s\rangle \models_F \phi \text{ und } \langle M, s\rangle \models_F \psi \\
\langle M, s\rangle \models_T \neg\phi & :\Leftrightarrow & \langle M, s\rangle \models_F \phi \\
\langle M, s\rangle \models_F \neg\phi & :\Leftrightarrow & \langle M, s\rangle \models_T \phi \\
\langle M, s\rangle \models_T B\phi & :\Leftrightarrow & \forall s' \in B.\langle M, s'\rangle \models_T \phi \\
\langle M, s\rangle \models_F B\phi & :\Leftrightarrow & \langle M, s\rangle \not\models_T B\phi \\
\langle M, s\rangle \models_T L\phi & :\Leftrightarrow & \forall s' \in W(B).\langle M, s'\rangle \models_T \phi \\
\langle M, s\rangle \models_F L\phi & :\Leftrightarrow & \langle M, s\rangle \not\models_T L\phi
\end{array}
$$

wobei für die Funktion $W\colon S \to 2^S$ gilt: Gegeben eine Situation s, so enthält $W(s)$ diejenigen Situationen, die mit s bzgl. der Wahrheitswerte der atomaren Aussagen übereinstimmen und zudem jeder atomaren Aussage genau einen Wahrheitswert ($true$ oder $false$) zuordnet. W wird wie folgt erweitert: $W(\{s_1, \ldots, s_n\}) := W(s_1) \cup \cdots \cup W(s_n)$.
Die Relationen $\models_T$ und $\models_F$ beschreiben somit, ob in einer Situation bzgl. des gegebenen Modells eine Aussage wahr oder falsch ist. Die ersten sechs Regeln bringen im Vergleich zu Definition 2.3 keine Neuerungen. Von besonderer Relevanz sind entsprechend die übrigen Regeln:
Die siebente Regel besagt, dass insbesondere Tautologien im Sinne von explizitem Wissen nicht in allen Situationen wahr sein müssen. Dieses bedeutet, dass keine Abgeschlossenheit des expliziten Wissens unter logischer Schlussfolgerung vorliegt und somit das Problem der logischen Allwissenheit diesbzgl. nicht auftritt. Die achte Regel besagt darüberhinaus, dass kein explizites Wissen bzgl. der Wahrheit einer Aussage vorliegt, falls diese in auch nur einer der dem Agenten zugeordneten Situationen nicht vorliegt (d.h. dieser Aussage der Wahrheitswert $false$ oder auch kein Wahrheitswert zugewiesen ist).
Für den Operator des impliziten Wissens werden nur die Situationen betrachtet, die einer Welt im Sinne von Definition 2.2 entsprechen. Er hat die Eigenschaften des Notwendigkeitsoperators (Wissensoperators) in einem S5-System. Darüber hinaus gilt aufgrund der Definition von W und den entsprechenden semantischen Regeln:

$$\models B\phi \Rightarrow L\phi$$

Kritikpunkte an dem von Levesque vorgeschlagenen Ansatz sind unter anderem, dass weder Quantifizierung noch Verschachtelung der Wissensoperatoren zugelassen sind. Zudem wiesen Fagin und Halpern nach, dass unter gewissen Umständen Agenten zwangsläufig Kenntnis aller aussagenlogischen Tautologien haben müssen.

3.1.2 Fagin, Halpern: Logic of general awareness

Die von Fagin und Halpern entwickelte Logik[2] enthält neben den Operatoren L für implizites Wissen und B für explizites Wissen einen Operator A für „sich bewußt/gewahr sein" (von engl.: Awareness; die zugehörige formale Sprache wird mit LGA bezeichnet). Ein Modell hat folgende Gestalt:

$$M = \langle S, A, R, \pi \rangle$$

wobei:

S	Zustandsmenge
$A : S \to 2^{Form(LGA)}$	Abbildung, die jedem Zustand diejenigen Formeln zuweist, derer sich der Agent in diesem bewußt ist
$R \subseteq S \times S$	transitive, serielle und euklidische (Erreichbarkeits-) Relation
$\pi : S \to 2^{At}$	zustandsabhängige Interpretation der atomaren Aussagen

Die semantischen Regeln aus Definition 2.3 behalten ihre Gültigkeit und werden durch folgende Regeln ergänzt:

$$\langle M, s\rangle \models L\phi \quad :\Leftrightarrow \quad \forall s' \in S \text{ mit } (s, s') \in R. \langle M, s'\rangle \models \phi$$
$$\langle M, s\rangle \models A\phi \quad :\Leftrightarrow \quad \phi \in A(s)$$

Im Gegesatz zu Levesques „Logic of implicit and explicit belief" wird der Operator des expliziten Wissens als Abkürzung eingeführt:

$$B\phi := L\phi \wedge A\phi$$

Das anhand des L-Operators definierte System entspricht aufgrund der Eigenschaften von R einem KD45-System (s. Kap. 2.1.3). Explizites Wissen liegt dann vor, falls dieses Wissen in allen erreichbaren Zuständen (Welten) vorliegt und sich der Agent des Wissens zudem bewußt ist.
Bzgl. der Diskussion dieses Ansatzes sei auf entsprechende Literatur verwiesen[3].

3.2 Agentenmodellierung: weitere Aspekte

Im Rahmen dieser Ausarbeitung erfolgte eine Betrachtung von ausgewählten logikbasierten Ansätzen zur Formalisierung des Wissens-/Glaubensbegriffs im Kontext der Agentenmodellierung.
Der Wissensbegriff ist allerdings nur ein Aspekt, der bei der Modellierung eines Agenten zu berücksichtigen ist. Andere Aspekte sind intentionale Begriffe wie Absichten, Wünsche und Ziele sowie deren Zusammenhang mit dem Wissen/den Informationen eines Agenten. Zu einer weitergehenden Diskussion des Begriffs der Intentionen sei auf **Bratman, M**, 1987: *Intentions, Plans, and Practical Reason* (Harvard University Press) verwiesen.

[2]siehe **Fagin, R., Halpern, J.Y.**, 1985: *Belief, awareness and limited reasoning*. In: IJCAI-85
[3]z.B. **Konolige, K.**, 1986: *A Deduction Model of Belief*. Pitman Publishing: London and Morgan Kaufmann.

Weiterhin bedarf die Agentenmodellierung der Möglichkeit, Dynamik ausdrücken zu können (wie verändert sich der kognitive Zustand eines Agenten über die Zeit? wie beeinflussen Umgebung und Informationen den Zustand des Agenten?), sowie die Möglichkeit, Pläne/Strategien zum Erreichen der Ziele des Agenten entwerfen zu können. Einen formalen Ansatz bieten Cohen und Levesque mit ihrer „logic of rational agency“[4], einer mehrsortigen und multimodalen Logik. Auch wird man in diesem Zusammenhang auf den Begriff der BDI-Architekturen treffen, hierzu sei z.B. auf **Georgeff, M.P., Rao, A.S.**, 1991: *Modelling Rational Agents BDI-Architecture* (in: Tech. Rep. 14, Australian Artificial Intelligence Institute, Melbourne) verwiesen.
Soziale Aspekte erlangen insbesondere in Multiagentensystemen mit einer großen Anzahl von Agenten Bedeutung. In diesem Kontext geht es u.a. um Fragestellungen der zweckmäßigen Interaktion von Agenten, um Koordination und Kooperation sowie den Begriff der sozialen Norm. Eine Diskussion dieser Begriffe ist z.B. **Jennings, N.R., Norman, T.J., Panzarasa, P.**, 2002: *Formalising collaborative decision making and practical reasoning in multi-agent systems* (in: Journal of Logic and Computation 12 (1), p. 55-117) zu entnehmen.
Eine Formalisierung des Wissensbegriffs und eine Verwendung dieser ist folglich nur ein Aspekt der Agentenmodellierung, welcher allerdings als Basis einer umfassenderen Modellierung verwendet werden kann bzw. sollte (man betrachte z.B. den intentionalen Begriff des „Ziels“, welcher wahrscheinlich Bestandteil jeder (nicht ausschließlich reaktiven) Agentmodellierung ist. Nimmt man o.B.d.A. an, dass das Ziel eines Agenten das Erreichen eines (bestimmten) Zustandes ist, so ist das Wissen über die Existenz dieses Zustandes eine notwendige Bedingung hierfür. Die Modellierung des Begriffs des „Ziels“ setzt folglich eine Modellierung des Wissens(-begriffs) voraus !).

[4] siehe **Cohen, P.R., Levesque, H.**, 1990: *Intention is choice with commitment.* Artificial Intelligence (42), p.213-261.

Literaturverzeichnis

[HABEL, ESCHENBACH (2003)] **Habel, C., Eschenbach, C.**, 2003:
Prüfungsunterlagen LOS — Logik & Semantik.
Fachbereich Informatik, Universität Hamburg.

[HALPERN, MOSES (1996)] **Halpern, J.Y., Moses, Y.**, 1996:
A guide to completeness and complexity for modal logics of knowledge and belief.
http://www.cs.cornell.edu/home/halpern/papers/guide.pdf. (01.11.2003)

[HUGHES (1978)] **Hughes, G.E., Cresswell, M.J.**, 1978:
Einführung in die Modallogik.
de Gruyter.

[LEVESQUE (2000)] **Levesque, H., Lakemeyer, G.**, 2000:
The Logic of Knowledge Bases.
MIT Press.

[NONNENGART, OHLBACH (1992)] **Nonnengart, A., Ohlbach, H.J.**, 1992.:
Modal- und Temporallogik.
In Bläsius, K.H., Bürckert, H.J., *Deduktionssysteme* (2. erweiterte Auflage, Kap.7).
Oldenbourg.

[POOLE, MACKWORTH, GOEBEL (1998)] **Poole, D., Mackworth, A., Goebel, R.**, 1998: *Computational Intelligence — A Logical Approach.*
Oxford University Press.

[SCHNEIDER (2002)] **Schneider, T.**, 2002:
Komplexität modaler Logiken.
Diplomarbeit, Jena.

[WIKIPEDIA] **Wikipedia**. http://de.wikipedia.org/wiki/Wissen. (24.03.2004)

[WOOLDRIDGE (1992)] **Wooldridge, M.J.**, 1992:
The Logical Modelling of Computational Multi-Agent Systems.
http://www.csc.liv.ac.uk/~mjw/pubs/thesis.pdf. (23.03.2004)

Abbildungsverzeichnis